REPORT DOCUMENTATION PAGE	Form Approved OMB NO. 0704-0188

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (From - To)
04-05-2010	Final Report	1-Aug-2006 - 31-Jul-2009

4. TITLE AND SUBTITLE
A Sequential Monte Carlo Method for Real-time Tracking of Multiple Targets: Final Report

5a. CONTRACT NUMBER
W911NF-06-1-0329

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
611102

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

6. AUTHORS
Scott T. Acton, Bing Li

7. PERFORMING ORGANIZATION NAMES AND ADDRESSES
University of Virginia
Office of Sponsored Programs
1001 N. Emmett St. P.O. Box 400195
Charlottesville, VA 22904 -4195

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSOR/MONITOR'S ACRONYM(S)
ARO

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
46850-CS.1

12. DISTRIBUTION AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES
The views, opinions and/or findings contained in this report are those of the author(s) and should not contrued as an official Department of the Army position, policy or decision, unless so designated by other documentation.

14. ABSTRACT
In this project, a Monte Carlo approach to tracking was developed for tracking in cluttered environments and across multiple scales. The Monte Carlo approach was compared with an active contour approach. Specifically, we developed a novel deterministic approach for 2D projective/affine snakes, which was evaluated in conditions of high clutter and with targets of varying viewpoint and scale.
A base view active contour method has been developed and tested for target tracking. The base view active contour

15. SUBJECT TERMS
Target tracking, image processing

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	15. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Scott Acton
UU	UU	UU	UU		19b. TELEPHONE NUMBER
					434-982-2003

Report Title

A Sequential Monte Carlo Method for Real-time Tracking of Multiple Targets: Final Report

ABSTRACT

In this project, a Monte Carlo approach to tracking was developed for tracking in cluttered environments and across multiple scales. The Monte Carlo approach was compared with an active contour approach. Specifically, we developed a novel deterministic approach for 2D projective/affine snakes, which was evaluated in conditions of high clutter and with targets of varying viewpoint and scale.

A base view active contour method has been developed and tested for target tracking. The base view active contour displayed an average error 10% more accurate than the correlation tracker and 14% more accurate than the centroid tracker tested with 120 synthetic videos corrupted with both Gaussian and impulse noise. Over 46 real video sequences base view active contours successfully tracked the target in an average of 80% of the frames as compared to 73% of the frames for the centroid tracker and 83% for the correlation tracker. When the real video sequences containing target occlusion were removed from consideration, the base view active contour successfully tracked in an average 87% of the frames whereas the correlation tracker's performance dropped to only 75% of the frames. Overall, base view active contours outperform the competing methods in the synthetic and real video experiments.

List of papers submitted or published that acknowledge ARO support during this reporting period. List the papers, including journal references, in the following categories:

(a) Papers published in peer-reviewed journals (N/A for none)

B. Li and S.T. Acton, "Automatic active model initialization via Poisson inverse gradient," IEEE Transactions on Image Processing, vol. 17, pp. 1406-1420, 2008.

A. Aksel, J.A. Hossack and S.T. Acton, "ACU-SEG: Automated cardiac ultrasound image segmentation and tracking," CSI Communications, vol. 32, pp. 11-16, 2008.

J. Cui, N. Ray, S.T. Acton and Z. Lin, "An affine invariance approach to cell tracking," Computerized Medical Imaging and Graphics, vol. 32, pp. 554-565, 2008.

P. Tay, B. Li, C.D. Garson, S.T. Acton and J.A. Hossack, "Left ventricle segmentation using model fitting and active surfaces," Journal of Signal Processing Systems, 2008.

B. Li and S.T. Acton, "Active contour external force using vector field convolution for image segmentation," IEEE Transactions on Image Processing, vol. 16, pp. 2096-2097, 2007.

D.P. Mukherjee and S.T. Acton, "Affine and projective active contour models," Pattern Recognition, vol. 40, pp. 920-930, 2007.

J. Cui, S.T. Acton and Z. Lin, "Tracking rolling leukocytes in vivo by factored sampling," Medical Image Analysis, vol. 10, pp. 598-610, 2006.

A.L. Klibanov, J.J. Rychak, W.C. Yang, B. Li, S.T. Acton, J.R. Lindner, K. Ley and S. Kaul, "A simple and efficient preparation of targeted ultrasound contrast agents for molecular imaging of inflammation in high-shear flow," Contrast Media and Molecular Imaging, vol. 1, pp. 259-266, 2006.

S. Millington, B. Li, J. Tang, S. Trattnig, J.R. Crandall, S.R. Hurwitz, S.T. Acton, "Quantitative and topographical evaluation of ankle articular cartilage using high resolution MRI," Journal of Orthopaedic Research, vol. 25, pp. 143-151, 2006.

J.J. Rychak, B. Li, S.T. Acton, A. Leppanen, R.D. Cummings, K. Ley, A.L. Klibanov, "Selectin ligands promote ultrasound contrast agent adhesion under shear flow," Molecular Pharmaceuticals, vol. 3, pp. 516-524, 2006.

J. Cui, S.T. Acton and Z. Lin, "Tracking rolling leukocytes in vivo by factored sampling," Medical Image Analysis, vol. 10, pp. 598-610, 2006.

G. Dong and S.T. Acton, "On the convergence of bilateral filter for edge-preserving image smoothing," IEEE Signal Processing Letters, 2006.

S. Millington, B. Li, J. Tang, S. Trattnig, J.R. Crandall, S.R. Hurwitz, S.T. Acton, "A quantitative study of human ankle articular cartilage and evaluation of ankle cartilage topography using high resolution cartilage sensitive MR imaging," Osteoarthritis and Cartilage, 2006.

S.T. Acton, "Snakes for tracking via generalized deterministic annealing," Journal of Electronic Imaging, vol. 14, 023017, pp. 1-13, 2005.

Number of Papers published in peer-reviewed journals: 13.00

(b) Papers published in non-peer-reviewed journals or in conference proceedings (N/A for none)

Number of Papers published in non peer-reviewed journals: 0.00

(c) Presentations

Number of Presentations: 0.00

Non Peer-Reviewed Conference Proceeding publications (other than abstracts):

Number of Non Peer-Reviewed Conference Proceeding publications (other than abstracts): 0

Peer-Reviewed Conference Proceeding publications (other than abstracts):

S. Basu and S.T. Acton, “Manifold learning for high content screening with automated cell segmentation,” Asilomar Conf. on Signals, Systems and Computers, Pacific Grove, California, November 4-7, 2007.

S. Basu, D.P. Mukherjee and S.T. Acton, “Implicit evolution of open-ended curves,” Proc. IEEE Int. Conf. on Image Processing, San Antonio, Texas, September 16-19, 2007.

B. Li, S.A. Millington, D.D. Anderson and S.T. Acton, “Registration of surfaces to 3D images using rigid body surfaces,” Asilomar Conf. on Signals, Systems and Computers, Pacific Grove, California, October 29 - November 1, 2006.

B. Li, A.V. Patil, J.A. Hossack and S.T. Acton, “3D Segmentation of the prostate via Poisson inverse gradient initialization,” Proc. IEEE Int. Conf. on Image Processing, San Antonio, Texas, September 16-19, 2007.

B. Li and S.T Acton, "Vector field convolution for image segmentation using snakes," Proc. IEEE Int. Conf. on Image Processing, Atlanta, Georgia, October 8-11, 2006.

B. Li and S.T. Acton, "Feature weighted active contours", Proc. IEEE Southwest Symposium on Image Analysis and Interpretation, Denver, Colorado, March 26-28, 2006.

B. Li, P. Tay and S.T. Acton, "Multi-assignment interacting multiple model for tracking microbubbles," Asilomar Conf. on Signals, Systems and Computers, Pacific Grove, California, October 30 - November 2, 2005.

Number of Peer-Reviewed Conference Proceeding publications (other than abstracts): 7

(d) Manuscripts

Number of Manuscripts: 0.00

Patents Submitted

Patents Awarded

Graduate Students

NAME	PERCENT_SUPPORTED
Alla Aksel	0.50
Bing Li	0.50
FTE Equivalent:	**1.00**
Total Number:	**2**

Names of Post Doctorates

NAME	PERCENT_SUPPORTED
FTE Equivalent:	
Total Number:	

Names of Faculty Supported

NAME	PERCENT_SUPPORTED	National Academy Member
Scott T. Acton	0.16	No
FTE Equivalent:	**0.16**	
Total Number:	**1**	

Names of Under Graduate students supported

NAME	PERCENT_SUPPORTED
FTE Equivalent:	
Total Number:	

Student Metrics

This section only applies to graduating undergraduates supported by this agreement in this reporting period

The number of undergraduates funded by this agreement who graduated during this period: 0.00

The number of undergraduates funded by this agreement who graduated during this period with a degree in science, mathematics, engineering, or technology fields:...... 0.00

The number of undergraduates funded by your agreement who graduated during this period and will continue to pursue a graduate or Ph.D. degree in science, mathematics, engineering, or technology fields:...... 0.00

Number of graduating undergraduates who achieved a 3.5 GPA to 4.0 (4.0 max scale):...... 0.00

Number of graduating undergraduates funded by a DoD funded Center of Excellence grant for Education, Research and Engineering:...... 0.00

The number of undergraduates funded by your agreement who graduated during this period and intend to work for the Department of Defense 0.00

The number of undergraduates funded by your agreement who graduated during this period and will receive scholarships or fellowships for further studies in science, mathematics, engineering or technology fields:...... 0.00

Names of Personnel receiving masters degrees

NAME	
Bing Li	
Alla Aksel	
Total Number:	**2**

Names of personnel receiving PHDs

NAME	
Bing Li	
Total Number:	**1**

Names of other research staff

NAME	PERCENT_SUPPORTED
FTE Equivalent:	
Total Number:	

Sub Contractors (DD882)

Inventions (DD882)

A Sequential Monte Carlo Method for Real-time Tracking of Multiple Targets

46850-CI
Scott T. Acton

University of Virginia

Scientific Progress and Accomplishments

A major stride has been made in terms of implementing and testing the base view active contour approach. Base view active contours are an improvement upon basic non-rigid active contours for tracking rigid vehicle targets. The base view active contour assumes that a vehicle's contour evolution can be represented by a finite set of contours. We look to exploit knowledge of the target vehicle to improve upon the non-rigid active contour in several ways.

The first goal is to increase the accuracy of the segmentation under noisy conditions. Base view active contours limit the way the active contour is allowed to evolve to only those shapes that the vehicle is known to be capable of appearing, thus eliminating spurious contour shapes resulting from noise in the system. Secondly, base view active contours seek to provide information about the pose of the vehicle of interest. Because of the predefined states which define the base views, knowing what state the algorithm is in with concern to segmentation and tracking also gives information about which way the vehicle is facing. This vehicle pose knowledge gives information that normally would require a human observer to interpret. One example of this is that knowing the pose of a vehicle as well as its direction of motion can tell the observer whether a vehicle is backing up as opposed to driving forward. The pose information could also be used in conjunction with a Kalman technique to provide refined observation in that the car should only have velocity relative to the direction it is facing.

A. Base View Sets

The idea behind the base view active contour begins with the intuitive notion that the more one knows about the shape of an object the better one should be able to track that object. The base views acquired for the proposed algorithm seek to describe the possible ways in which a camera might image the vehicle of interest. Base views encompass as wide a range of angles of the vehicle as possible through the evenly spaced sampling of the vehicle's image space. This image space, as defined by this algorithm, includes all the possible 2D images that could be acquired by rotating a camera about a 3D model of the vehicle of interest, where the camera distance from the vehicle is fixed at some arbitrary distance such that, at any angle, the entire vehicle is visible and not cut out of the camera's field of view.

There are two important points in this definition of the base view sets. One is that the sets must be sampled evenly throughout the image space. After more preprocessing of the base view sets, this provides even spacing of the base view active contour states. Even spacing of the base view sets is important for the accuracy of the tracking and segmentation as well as having a well behaved transition between states. The behavior of states that are too far apart relative to the other states in its neighborhood presents a decrease in the likelihood for transition to that particular state which may or may not be a desired behavior.

The second important point for the definition of the base view sets is that the camera must be at a sufficient distance relative to the 3D model such that the entire vehicle can be seen. This stipulation requires all the of the base views to be acquired at the same relative scale, an important point for implementation, and that the whole vehicle

is in the field of view to permit a complete and accurate contour acquisition in the next phase of preprocessing.

Assuming that the algorithm would be applied to a surveillance application narrows down the area of the image space needed to sample to accurately represent the vehicle. In a street level surveillance application there are angles relative to the vehicle that the camera will not achieve, thus eliminating the need to sample those. For example, we need no images of the bottom of a vehicle because those perspectives would never be imaged in a surveillance situation. A less obvious application of this idea is that for a given camera field of view, one could reduce the size of the base view set to only those vehicle perspectives the camera captures. A camera looking out toward a parking lot would not necessarily require contours representing a direct top down view of a vehicle because that particular angle would not be obtained in its normal field of view.

B. Acquisition of Base View Sets

Since the quality of the of the base view sets has a direct impact on the performance of the algorithm great care was taken in their acquisition and preparation for use in the algorithm. The base views had to be taken at known and evenly spaced angles to assure thorough coverage of the image space of the vehicle. The lighting and background in the base views also had to be consistent; such conditions would allow the active contour method used to acquire the base view active contours more consistently and accurately represent the vehicle. In this way we seek to minimize the noise in the process of acquiring the base view contours of interest.

The 3D models used in capturing the base views are standard 1:18 scale. This model scale proved easy to image because of its convenient size and yet contained enough detail to accurately represent its respective vehicle. An electrically actuated, fully translating mechanical arm was utilized to ensure equal spacing of the acquired images. The mechanical arm system was capable of angles from 0 – 90 degrees in the φ direction and 0 – 360 degrees in the θ direction. As stated before, the distance between the camera and the vehicle, ρ, was held constant to fix the scale of the base view contours.

Utilizing the mechanical arm and a digital camera, images of the vehicle models were collected at evenly spaced intervals. The movements of the mechanical system are precise to the degree, though the base view sets collected were, at their finest, taken at tens of degrees. Taking the sets at tens of degrees avoids the mechanical error of taking sets near the precision of the mechanical system and also decreases the storage space necessary to contain the full base view sets. Given a more precise imaging system as well as no limits on storage of images the base view sets could be taken at any arbitrary interval of degrees. The smaller the sampling interval the closer each base view becomes to its neighbors in that the neighboring images will more closely resemble each other. Neighboring images are those images that are close to each other in the image space. The "image space" of an object is simply a term referring to the many potential perspective of an object that can be imaged by a camera. An example of this is shown in Figure 1.

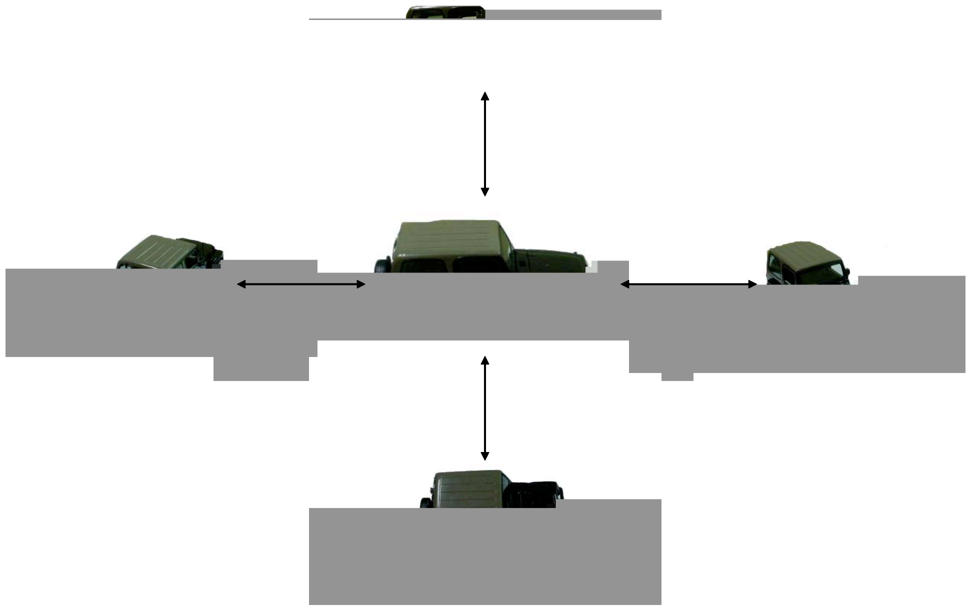

Figure 1. A neighborhood of vehicle images. The central image is the center of the neighborhood. Each of the outlying images is a neighbor to the central image because when the images were taken by the mechanical arm they were separated only by one sampling interval either in the φ or θ direction.

The primary reason, however, for taking coarse samples of the vehicles' image space is to demonstrate an assumption of this approach that a vehicle being imaged by a fixed camera will pass through distinct states that can be prerecorded. By taking a coarse sample of the image space, the images are far enough apart to be distinct from each other allowing the desired state based view of the movement of the vehicle through an image sequence. The base view active contour process is demonstrated in Figures 2-5.

Figure 2. A raw set of base view images.

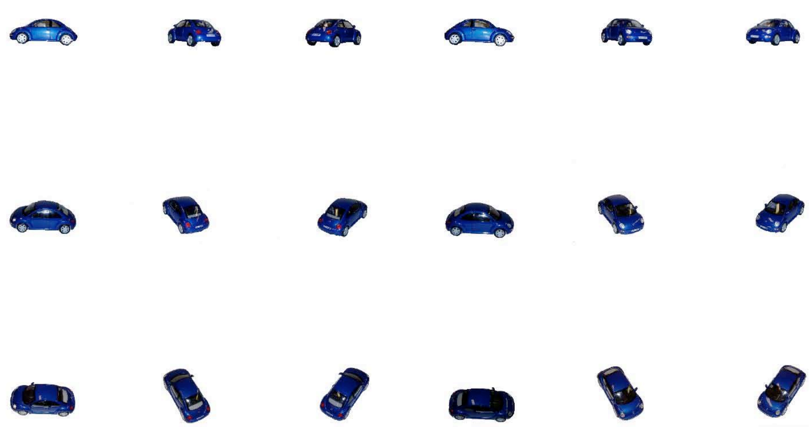

Figure 3. A base view set after background subtraction.

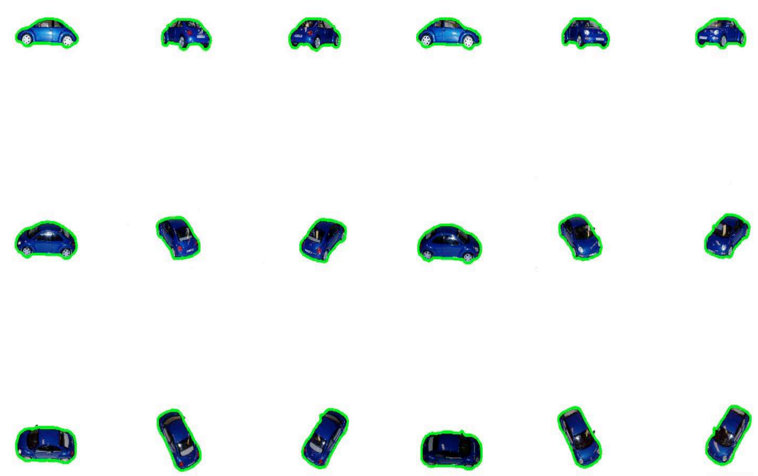

Figure 4. The base view set after a VFC snake find the contour around each view.

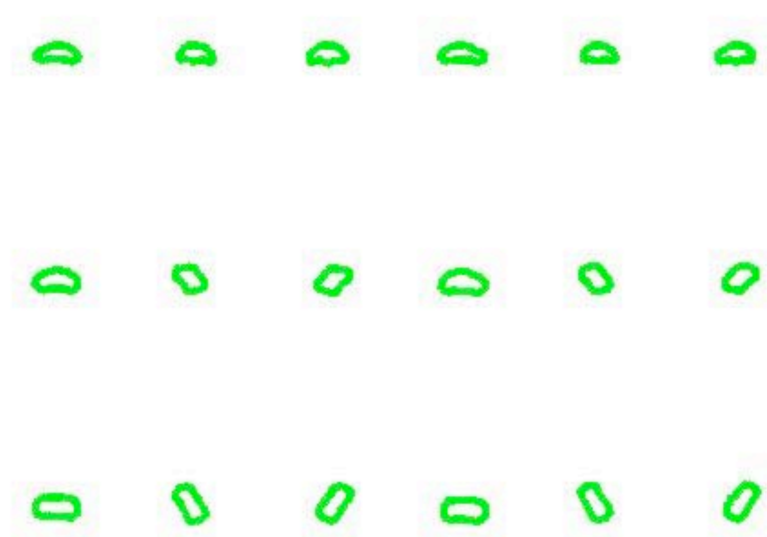

Figure 5. The final product of the preprocessing is the vehicle represented only by its shape at different views.

C. Preprocessing of Base View Sets

Since base view active contours utilize the contour of a vehicle and not a raw image template as used by a correlation tracker some preprocessing of the base views is required before their application to the tracking method. The first step in the preprocessing is to remove everything from the acquired base views that is not the vehicle in question. This is done through a simple background removal process that is made even simpler because of the preparations made during image acquisition. During image acquisition lighting and background were controlled so that the background would be a much different color than that of the vehicle itself. Using color segmentation, where the color of the background is sampled and averaged, the background is located and then removed from the image. The background was replaced by a flat white color. The choice of a uniformly white background provides a high level of contrast between the vehicle and the new background which improves the performance of the next step of the preprocessing, contour acquisition.

The contour acquisition process occurs after the full set of base views have been set against the uniformly white background. Then, view by view, a VFC active contour is applied to the base view set. Because this process is not limited by time constraints, noise, occlusions, or any other problems associated with tracking the active contours are given all the benefits necessary to ensure that they accurately capture the edges of the vehicle in each view.

The final step in preprocessing involves the preparations necessary for the use of the now base view active contours in the actual tracking algorithm. These preparations include the calculation of the area enclosed by the active contour as well as the center of

mass of each active contour. These two measures are necessary because though the contours themselves are now generalized, the description of the contours is based on the snaxel locations in these base view images. Having the initial areas and centers of mass located allows the contours to be easily translated and scaled during the tracking scenario. The contours, with their respective areas and centers of mass, are then stored so that they can be easily accessed and referenced by the tracking portion of the algorithm.

D. Base View Contours Applied to a Hidden Markov Model Architecture

Once the base views are collected a method is necessary to allow for accurate and quick transistions between them so that the best base view contour is used to track the vehicle in a given frame of video. A hidden Markov approach provides the architecture and underlying model for the base view active contour algorithm. The HMM is useful in this case because we wish to represent the change in shape of a vehicle throughout a tracking sequence as a transition between states which are a part of the vehicles shape space. The HMM accommodates this assumption because as discussed previously, the HMM assumes that the system being modeled is a Markov chain; that the next state depends only on the previous state. This makes intuitive sense in that the next shape of the vehicle in question should depend only on the shape of the vehicle in the previous state (or time step).

The base view contours that are collected for a particular vehicle provide one part of the HMM architecture immediately: the number of states N. N is defined simply by the number of base views taken because each view was taken to represent a state that the shape of the vehicle's outline might take during a tracking sequence. Other parts of the

HMM architecture can be taken indirectly from the base view contours. The state transition model is simply an organized version of the base view contours such that the "neighboring" states are no more than one step a part. That is, a neighborhood of base view contours consists of those contours whose similarity to each other comes from the fact that they were acquired at similar (closely related) angles of θ and φ. It follows that turning the vehicle left or right, up or down from the current state should transition to one of the neighboring states. Figure 6 shows an example Markov chain for a base view contour set. Note: The arrows at the edges of the figure do not necessarily transition to the state on the opposite edge of the figure.

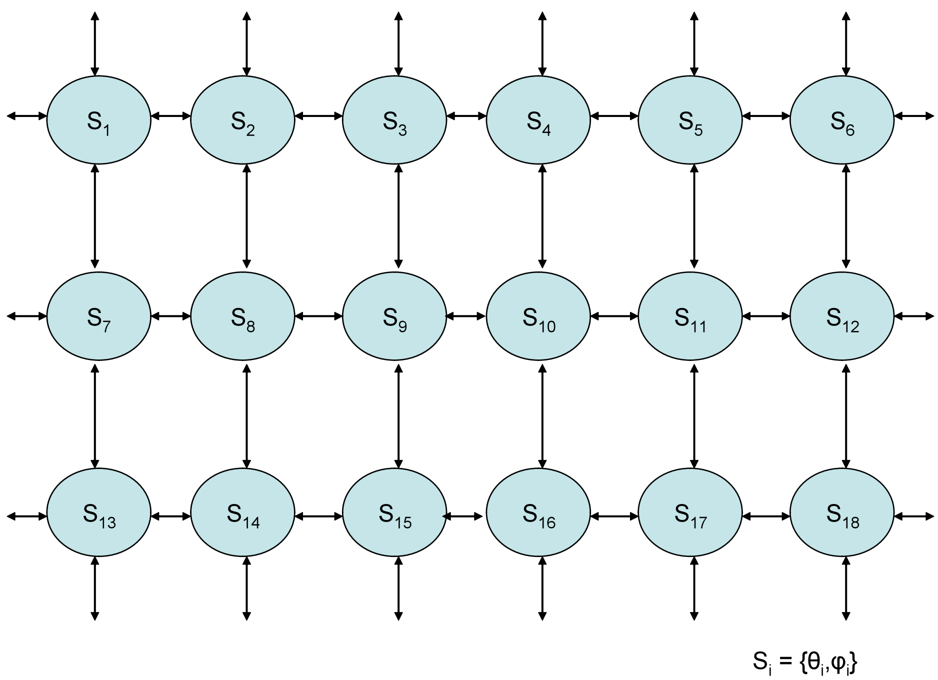

Figure 6. The Markov chain representing an N = 18 base view contour set. Each state is indexed by the angles of θ and φ that they were acquired at to simplify setting up the structure. Transitions are allowed between horizontal and vertical neighbors.

The Markov chain in Figure 6 also lends insight into the observation model number, *M*, for the HMM architecture. Looking at the neighborhood around a single state shows the possible observations the system should expect in the next time step. For example, in Figure 6, being in state S_9 at time *t*, the Markov property holds that at time *t+1* the state can only have progressed to states S_3, S_{10}, S_{15} or S_8. Therefore, the available observations are those representing the current state, S_9 , and its four neighboring states. This is true for every state in this Markov chain which gives an observation number, *M*, of five for each state. Figure 7 shows the observations neighborhoods for a portion of the states and how they may overlap.

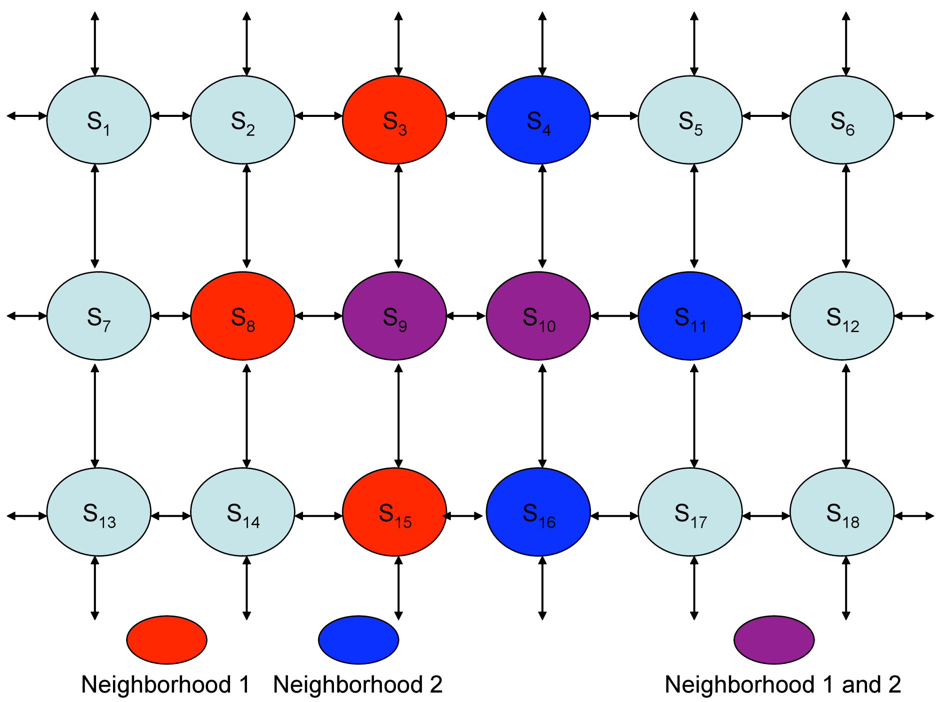

Figure 7. Observation neighborhoods within the hidden Markov model. States can be part of several different neighborhoods depending on what state the system is in at a given instant.

The state transition probability distribution, *A*, indicates to the model the states the HMM is allowed to transition to and from. Again, the Markov chain defined by the base view contours gives intuition in how this is formed. At first glance it appears that each state should have an equal likelihood of transition; the transition probability distribution should be uniform over a single neighborhood of five contours. Making the distribution uniform, however, creates an undesirable behavior. In some instances, where the vehicle being tracked is not changing its pose toward the camera (i.e. the state should not be transitioning when the vehicle is driving in a straight line across the field of view) the HMM would change states back and forth between multiple states within a neighborhood.

In order to smooth out this behavior, the probability density weighting was changed from that of a uniform shape to a more discretely Gaussian shape, i.e. the current state, being at the center of the Gaussian, is favored over the outlying states. In this way, the HMM is more likely to remain in the current state than make random transitions to neighboring states. In addition, to make the model more responsive to changes within the image, another weighting factor is added to the transition probability distribution. As we know from the discussion of active contours previously, an active contour seeks to minimize the external energy because it is closely related to the edges in the image. Therefore, we added a weighting based on the normalized external energy of the base view contour.

E. Base View Active Contours as Constrained VFC Active Contours

The active contour method put forth in this approach evolves the shape of the contours through the use of the HMM paradigm. Base view active contours remain

active contours in that they still seek to minimize the external energy by moving toward edges of interest. Base view active contours, however, have a rigid shape and therefore do not need to account for the internal energy of the contour explicitly. Rather, in between each step where the state of the HMM is changed, the base view contour of the current state is evolved through scaling and translation to better locate itself upon the edges of the target vehicle.

F. Vector Field Convolution and Base View Active Contours

Vector field convolution snakes have been shown in previous work to have large capture ranges while maintaining relative insensitivity to noise (Li and Acton, 2007). These favorable properties come from the way that the external force is calculated; the edge map of the image is convolved with a vector field pointing toward the origin. The resulting external force has vectors that point toward the nearest significant edge. Even though the evolution of base view active contours is different from a normal snake, the VFC external force characteristics are still desirable and the external force for base view active contours is constructed as described in (Li and Acton, 2007).

The external force, however, is where the similarity ends. The base view contours must be evolved in such a way that their shape is maintained and thereby remaining in the current state of the HMM. Thus, the only evolutions of the base view contours are scaling and translation because these do not alter the shape of the base view contours.

G. Results Summary

A base view active contour method has been developed and tested for target tracking. The base view active contour displayed an average error 10% more accurate than the correlation tracker and 14% more accurate than the centroid tracker tested with 120 synthetic videos corrupted with both Gaussian and impulse noise. Over 46 real video sequences base view active contours successfully tracked the target in an average of 80% of the frames as compared to 73% of the frames for the centroid tracker and 83% for the correlation tracker. When the real video sequences containing target occlusion were removed from consideration, the base view active contour successfully tracked in an average 87% of the frames whereas the correlation tracker's performance dropped to only 75% of the frames. None of the tracking methods tested in this work were designed to track under occlusion so removing real videos containing an occluded target gives a clearer indicator of the true relative performance of the trackers. Overall, base view active contours outperform the competing methods in the synthetic and real video experiments.

The PI and students at the University of Virginia has published these results in top journals and in major international conferences. A list of these publications is attached.